# Reflections on Regional Science in Business

**Graham CLARKE**
School of Geography, University of Leeds
Leeds LS2 9JT
UK
graham@geography.leeds.ac.uk

**Resumo**

*O objective deste artigo é o de resumir os argumentos de suporte à necessidade e vantagens da investigação aplicada e dar uma série de exemplos de aplicações práticas desenvolvidas na Universidade de Leeds.*

**Palavras-chave:** Ciência Regional, Sistemas de Informação Geográfica

**Abstract**

*The aim of this paper is to summarise the arguments pertaining to the need and benefits of applied research and then to give a series of examples of applied work undertaken at the University of Leeds.*

**Keywords:** Regional Science, Geographic Information Systems

**Résume**

*L'objectif de cet article est de faire un sommaire des arguments de support a la nécessité et avantages de la recherche appliqué et de présenter une série d'exemples d'applications pratiques développées a la Université de Leeds.*

**Mots clés:** Science Régional, Systèmes d'Information Géographique

## 1. Introduction

In a recent book (Clarke and Madden 2001) we review progress in applied regional science, especially as it relates to the interests of business and governments. The aim of this paper is to summarise the arguments pertaining to the need and benefits of applied research and then to give a series of examples of applied work undertaken at the University of Leeds. These examples can be read alongside the many non-University of Leeds examples given in the book.

A key paper on the development (and future) of regional science was published by Bailly and Coffey in 1994. They questioned the relevance of much of the work undertaken, not in terms of theoretical importance, but in terms of its usefulness to public and private sector planners. They talked of 'regional science in crisis' and encouraged regional scientists (largely geographers, economists and planners) to address real world problems and to return to basic concerns with people's everyday lives. The paper generated much debate in the regional science journals. Not all agreed of course. Many felt that discipline was alive and well and in no need of reform. However, it is clear that many researchers have got more involved with these issues since 1994. The book by Clarke and Madden (2001) was an attempt to put together a collection of papers from eminent regional scientists who had looked to apply their models. The criterion for selection was a client from business or government who had sponsored the work directly (i.e. not a normal research grant from a Government where individuals are eligible to apply for support for any type of research area). The book contains 17 chapters drawing upon applications in many key areas of regional science techniques, including input-output models, CGE models, spatial interaction models, land-use/transport models, dynamic models, population forecasting, optimisation techniques, microsimulation and GIS.

A key question is why is it now so much more common to see publications on applied spatial analysis than it was ten or twenty years ago? Perhaps many have simply followed the call to arms by Bailly and Coffey, as individual researchers enjoy the feeling associated with their methods being perceived as useful by the outside world. Certainly, there is a growing realisation that business and Governments can provide much needed new incomes for universities, as many traditional Government socio-economic research budgets decline. With the arrival of powerful PCs many businesspersons are better equipped to handle GIS or 'decision support systems' that academics can provide. Better, and certainly more routinely available data sets, also help to promote collaboration between academia and business. Lastly, we may argue that the benefits to teaching are also immense. Students find traditional lectures based on statistics and modelling rather dull! – however, if they can be spiced up through applications (in interesting areas such as retail location, airport location, social policy, trade flows, human migrations etc) then the students can more readily see the worth of such methodologies and may even find jobs where they are routinely applied (for example, companies such as GMAP and Eurodirect in Leeds have 75% of their workforce – 120 persons – recruited from the School of Geography at the University of Leeds).

In the rest of this article, a number of case study applications are presented. These are all taken from current research at the School of geography, University of Leeds. The two main methodologies explored are spatial interaction models and microsimulation models. The former can be used in a number of ways, from static models used for what-if analysis to models designed to find optimal network distributions. Examples of each type of model will be presented. As mentioned above, the reader is also encouraged to look at the chapters in the Clarke and Madden book (and see also Stillwell and Clarke 2003).

## 2. Applications of models in business: i) spatial interaction models

### 2.1. Introduction

Spatial interaction models are based on over thirty years of academic model development beginning principally with Wilson's derivation of the model through entropy maximisation (Wilson 1970, 1974). These models have provided a rich source of theoretical developments in understanding urban spatial structures. Applications of the models became more widespread from the mid 1980s onwards. For that reason, the models have also been incorporated into a number of proprietary GIS packages. However, applied work demonstrates that spatial modelling is highly context specific. There is no simple or standard approach, which has implications for the development of 'off-the-shelf' products. As is discussed in more detail elsewhere (Benoit and Clarke 1997) proprietary GIS packages provide very simple aggregate models with only one or two parameters. These can be of little practical use in dealing with the complexity apparent in most applied situations. It is hoped to demonstrate some of this complexity in the remainder of the paper.

The aggregate model can be written as:

$$S_{ij} = A_i e_i p_i W_j^{\alpha} e^{-\beta c_{ij}} \tag{8.1}$$

where

$$A_i = 1 / \sum_k W_k e^{-\beta c_{ik}} \tag{8.2}$$

to ensure that

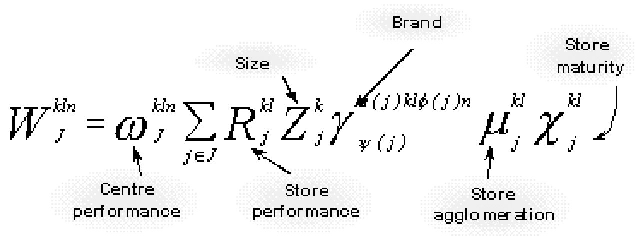
$$\sum_j S_{ij} = e_i P_i \tag{8.3}$$

The model can be used to estimate any type of flows in a city or region. For example, in a retailing context $S_{ij}$ is the flow of expenditure from residential zone *i* to shops (shopping centres or a store of a particular kind) in *j*; $e_i$ is the per capita expenditure of residents of *i*; $P_i$, the population of *i*; $W_j$, a measure of the 'attractiveness' of *j*, usually measured via floorspace as a proxy; $c_{ij}$, a measure of travel distance or cost between *i* and *j*. $\alpha$ and $\beta$ are parameters.

These models have been adapted to a wide range of application areas. They have been used to examine flows of people to shops, offices, work, schools, hospitals and even pubs and dry ski slopes! Indeed, Fischer and Getis (1999) remark that models of spatial interaction have been fundamental in regional science. The clients of the University of Leeds and GMAP Ltd (a company which arose out of the School) include ASDA, Sainsburys, Esso, Toyota, Ford, Halifax, Barclays, Thorn EMI, Storehouse and W.H.Smith. Although there are slight variations in what these individual organisations want from this collaboration, we can generally group these requirements together. Thus for banks, oil companies, car manufacturers and so on channel management and network planning are key strategic and operational issues.

The literature has examined many alternative versions of the variables in equations 1 to 3. As noted above, such disaggregations are crucial for models to work effectively in applied contexts (see Pacione 1974, Spencer 1978, Fotheringham 1986). For example, Eyre et al (2000) investigated the model results obtained within Yorkshire for a major UK high street retailer. Using conventional attractiveness variables, the models seriously under-predict the regional shopping centres such as Meadowhall and more 'attractive' centres such as Leeds and York. At the same time the models over-predict the performance of older, more run-down centres such as

Bradford. New variables were required that captured more about store and centre performance (for example, the fact that the store may be within a highly attractive, pedestrianised centre with many other major retailers present).




These kinds of study show that the supply-side variable is more likely, in reality, to take the following form:

Again, this demonstrates the difficulty of using very simple aggregate models in most applied situations (cf. Benoit and Clarke 1997).

The ultimate success of this sort of modelling lies in its predictive power. In most markets the models need to be within 10% of real turnovers (when known) 80% of the time. If actual customer flow data is available, then the models can produce even better levels of performance. Once built, the key use of the models is to predict the impacts of changes (demand or supply side changes). This will be explored below.

**2.2 Comparative static mode**

Once the models have been calibrated they can be used for analysis purposes. A typical product of the flow model for retailing is revenue, Dj:

$$D_j = \sum_i S_{ij} \tag{8.4}$$

This is not only useful for calibration purposes. The analyst is able to examine predicted turnover versus actual turnover (thus producing a measure of store performance) and to use the models to estimate the turnover of all competitor stores (which they are unlikely to know). Another important indicator produced by the models is small-area market shares. Most organisations know their national market share, and many will be able to disaggregate this to the regional level (and there are often market intelligence reports to help here). However, very few know their market shares at the local level (Birkin et al 1996). Mapping local market shares provides an immediate snapshot of retail performance. It shows areas where the organisation has few customers. These spatial gaps in the market become possible new store locations for any firm that is contemplating expansion. Having understood the existing performance of individual branches or outlets the models are then used in ‘what-if?’ fashion. Each one of the variables given in the models can be modified and the models re-run to estimate the impacts.

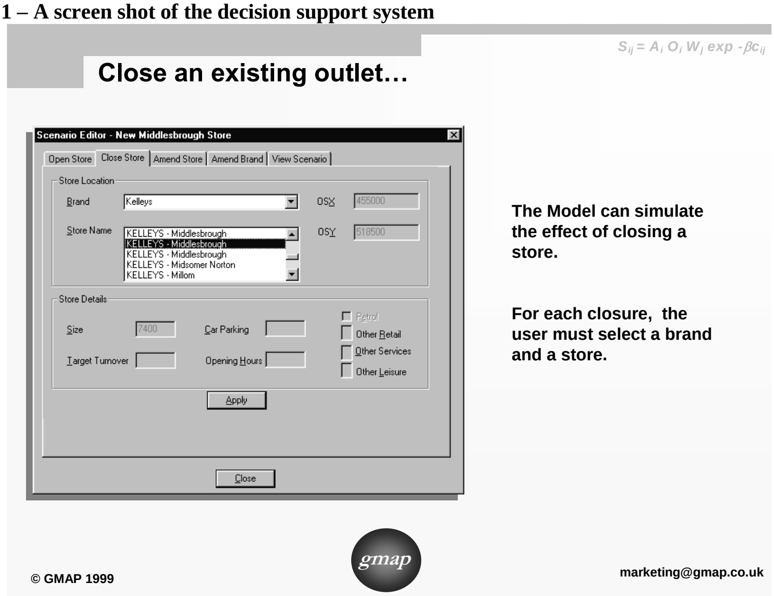
The illustration which follows is based on a major car company in the UK. The effects of simple or complex network changes may be evaluated, for example opening of new outlets, closure of existing outlets, outlet reinvestment and outlet rebranding. The model operates by simulating how customer patterns change for the network reconfigurations. The impacts upon sales and market share are evaluated and this provides valuable information to the investment appraisal process. The business model allows the impacts of various scenarios to be viewed: by outlet, by geographical area, by brand and by product line. Table 1 shows the results for a new car dealer opened in Blackpool in North West England. The model not only predicts the new sales total (310) but also shows how many sales will be lost (or deflected) from the company's existing outlets in the region. In this instance, it can be seen that the loss of sales within the company is low overall, although the nearby St Annes dealer may lose around 20 sales per annum (table 1).

It is important to package these models within an easy to use interface. Fig 1 shows the screen image of one such 'decision support system'. Hierarchical menu systems allow the user to be guided through the modelling options. In the case of the example shown in Fig 1, the user has selected the 'close store' option. More examples are provided in Clarke and Clarke (2001).

**Table 1 – Impacts of opening a new car dealership in Blackpool on the retailer's own outlets**

| Outlets | previous sales | new sales ($000 per week) |
|---|---|---|
| New Outlet in Blackpool | 0 | 310 |
| Chorley | 448 | 445 |
| Preston | 805 | 796 |
| St.Annes | 412 | 391 |

**Figure 1 – A screen shot of the decision support system**




## 2.3. Dynamic mode

A significant advance in spatial interaction modelling was the possibility of modelling the *evolution* of urban structures. This was first proposed by Harris and Wilson (1978). Taking the model presented by equations (1)-(3) the next step is to define retailers' costs, Cj, which can be taken as a function of a lead variable such as floorspace, Wj:

$$C_j = C_j(W_j) \qquad (4)$$

or more specifically for a simple illustration,

$$C_j = k_j W_j \qquad (5)$$

A typical product of the flow model is revenue, Dj:

$$D_j = \sum_i S_{ij} \qquad (6)$$

Thus, profit, $\prod_j$ at j is

$$\prod_j = D_j - C_j \qquad (7)$$

$$= D_j - k_j W_j \tag{8}$$

However, if 'normal' profits are included in retailers' costs, then the market will ensure that the $\prod_j s$ are competed to zero and so the following equilibrium condition must hold:

$$D_j = k_j W_j \tag{9}$$

In full, substituting for Dj from (6) and Sij from (1) and (2), this can be written

$$\sum_i \frac{e_i p_i W_j^{\alpha} e^{-\beta c_{ij}}}{\sum_k W_k^{\alpha} e^{-\beta c_{ik}}} = k_j W_j \tag{10}$$

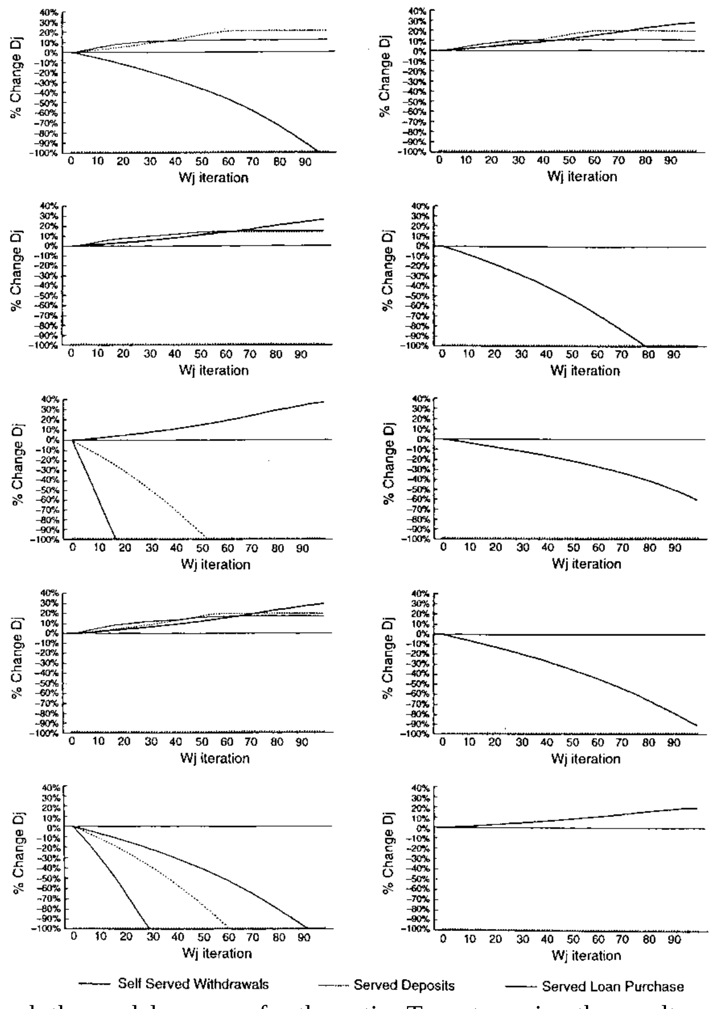
which are a set of non-linear simultaneous equations in the Wjs.

Thus, if revenue exceeds costs then a particular retail centre is likely to grow. Conversely, if costs exceed revenue then the location of the facility is unstable and the centre will decline. This process continues in an iterative manner until the solutions reach stability and no further growth/decline takes place unless further change is introduced to any of the model's variables or parameters. When compared to existing structures the equilibrium model must be interpreted carefully. It could be argued that the models are more concerned with the nature of equilibrium solutions (initial and changed or final solutions) than the process of dynamics itself. However, in moving from initial to final solutions there needs to be some consideration of the key processes of change which the variables and parameters of the models are trying to capture.

The main question to be addressed in this section is can the dynamic model play a useful role in contemporary store location research? The context for this real-world application comes from the desire of a major Canadian bank to reduce the number of branches in Toronto as part of a strategy to cut costs in the more competitive financial service market of the mid 1990s. The problem with branch closures per se however is that there is a strong correlation between new customer attraction and bank locations (despite electronic banking etc.: see Birkin et al 1996). So, the question is can we devise a strategy whereby branches may not necessarily be closed, but costs are reduced by removing the number of facilities offered at each branch. In other words can we *reconfigure* the products and services available at each branch according to local consumer demands.

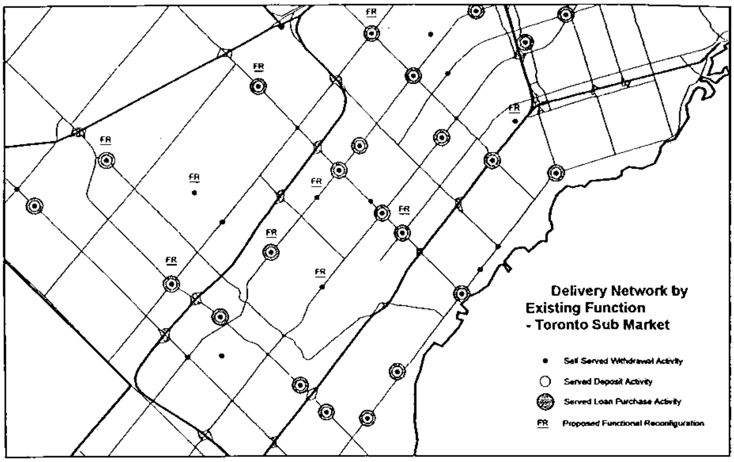
The basic objective of exploring the Dj=Wj equilibrium mechanism, therefore, was to identify the most unstable components of the delivery network, from a consumer usage perspective, given controlled changes to the calibrated model. The Dj=Wj equilibrium mechanism allowed for the identification and assessment of poor performing outlets in respect to stability (in fact the application showed that a revised Dj=Wj mechanism was required – see Clarke et al 1998). Those delivery outlets identified as most unstable within the sub set of poor performers became the closure candidates. It should be noted that a candidate for closure relates to the removal of a specific activity from the delivery outlet and not necessarily the absolute closure of the outlet.

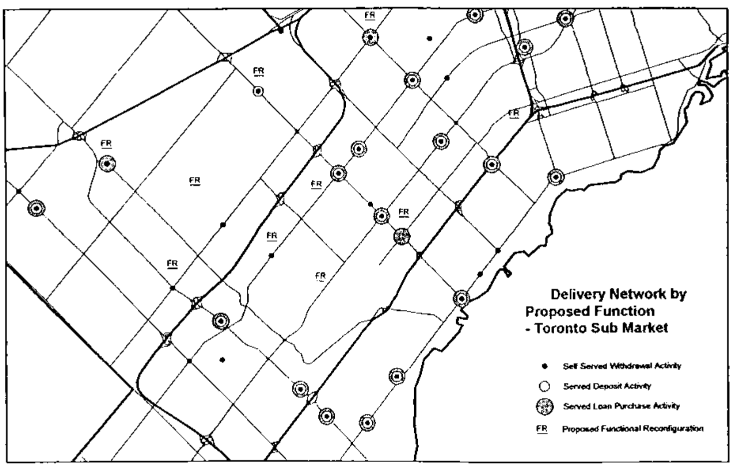
**Figure 2 – Dynamic Mode**




Although the model was run for the entire Toronto region the results we present here are based on a small sub-market. These patterns of network stability by activity type are presented in Figure 2. The graphical representation of the rate of change in Dj by activity type allows for an easy comparison of network stability between and within the branches of the sub-market of Toronto. Three product lines are modelled: the self served withdrawal, served deposit and loan purchase activity. Upward lines on the graph indicate stable and profitable sites for those products. Lines which decline towards zero on the Y-axis show unstable locations and potential candidates for closure (by product). The mapped version of the local reconfiguration plan (Fig 3) demonstrates the specific and localised nature of the plan.

**Figure 3 – Local reconfiguration plan**




The full financial benefits of this procedure are outlined in Clarke et al (1998). However, to summarise, it is estimated that this plan would save $13.4 million in operating costs (approximately 9% of the total), particularly through branch staff reduction.

## 2.4. Optimisation

So far the models have been used to open or close single stores, although the dynamic analysis in Toronto was undertaken for all stores simultaneously. A more radical question to ask is what an ideal distribution of facilities might look like. This is a relevant question to ask when major network restructuring is planned, or when new markets are under development (e.g. many western European service providers now entering Eastern Europe for the first time). The optimisation might be applied to a single region or to an entire country. The rest of the argument here is taken from Birkin and Culf (2001).

One solution procedure has now been developed using the concept of a ‘genetic algorithm’. Suppose that the problem is to select the best 50 locations from a universe of 8000 possible locations. This solution can be represented as a binary string of 8000 digits representing the possible locations. 50 of the digits are set randomly to 1, representing a starting point for the algorithms; the remainder are set to zero - these locations have not been selected. The ‘fitness’ of each solution is the number of cars which can be sold when dealers are placed at the specified locations (which are in turn calculated by summing the interactions for each dealership). The first step in the procedure is to create a population of random solutions, in which solution takes the form of an 8000 bit string, as described above. At each step of the algorithm, new solutions are created by selecting a pair of solutions from the initial population and combining them to form two new solutions. At regular intervals within the process, a ‘mutation’ may be introduced whereby any dealer is moved at random to a new location.

The process by which solution pairs are selected from the initial population is determined by a Monte Carlo selection process according to the fitness of the parent solution. At first there is little differentiation between solutions, but the selection process is made more aggressive as the solution develops, so that the characteristics of the fitter parents proliferate with time. The algorithm can be terminated after a fixed length of time (number of iterations) or when the change between iterations falls below a certain level. The performance of the genetic algorithm is extremely impressive. In all cases, the best solution is found, as compared with more ‘brute force’ methods.

It is with a certain amount of confidence, therefore, that the GA technique can be applied to ‘real world’ optimisation problems. In this section, the outputs from the procedure are demonstrated for a variety of car dealer networks in Denmark. Table 2 shows the existing dealer structure and network performance for six different manufacturers - Ford, Audi-VW, GM/Opel, Mercedes, Rover and Nissan. The first challenge for the algorithm is to relocate the existing dealers to their ideal locations. In the case of Ford, 105 dealers can be relocated in such a way that average sales are increased from 158 to 202 vehicles, an increase of 27.7%. It would appear that the most efficiently configured is that of Nissan (and it is probably no coincidence that this is the most sparse of the four networks). Even in this case, however, an average sales increase of more than 18% can be achieved through idealised relocation.

**Table 2 – Full Network IRP Run**

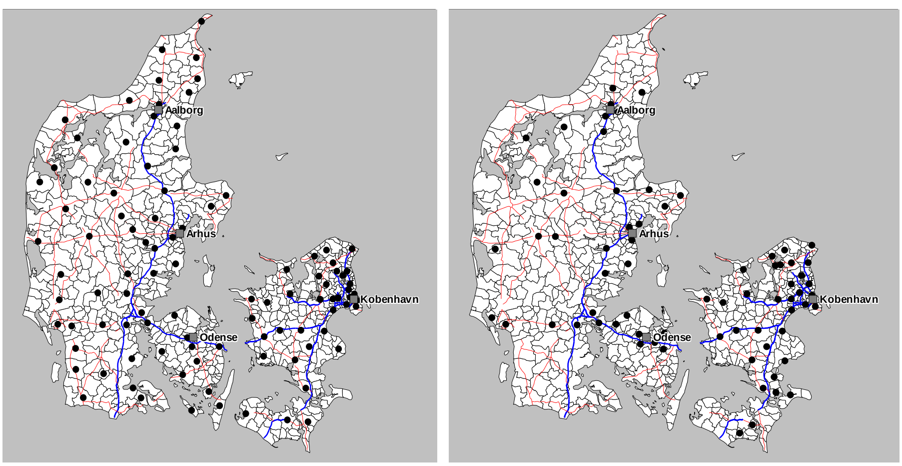
| | Existing Dealers | | | Full network | | | Change | | | % Change | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Manufacturer | Sales | Dealers | Average | Sales | Dealers | Average | Sales | Dealers | Average | Sales | Dealers | Average |
| Ford (01) | 16.615 | 105 | 158 | 21.217 | 105 | 202 | 4.601 | 0 | 44 | 27,7 | 0,0 | 27,7 |
| Audi-VW (03) | 19.745 | 101 | 195 | 26.024 | 101 | 258 | 6.279 | 0 | 62 | 31,8 | 0,0 | 31,8 |
| GM/Opel (09) | 17.590 | 100 | 176 | 23.371 | 100 | 234 | 5.782 | 0 | 58 | 32,9 | 0,0 | 32,9 |
| Mercedes (17) | 697 | 45 | 15 | 1.543 | 45 | 34 | 846 | 0 | 19 | 121,3 | 0,0 | 121,3 |
| Nissan (19) | 6.152 | 84 | 73 | 7.278 | 84 | 87 | 1.126 | 0 | 13 | 18,3 | 0,0 | 18,3 |
| Rover (23) | 1.275 | 23 | 55 | 3.736 | 23 | 162 | 2.461 | 0 | 107 | 192,9 | 0,0 | 192,9 |

The second challenge for the algorithm is to strip each network of dealers, so that the same number of cars can by sold through a restricted network of ideally located outlets. The results from these experiments are shown in Table 3. This can be achieved by a reduction in dealer numbers of between one-third and two-fifths the original quota for the various manufacturers. A comparison of the existing Ford network with the rationalised network is shown as Figure 4. It is clear, that such savings justify the costs of the models many times over! For more details of this application see Birkin and Culf (2001).

**Table 3 – Rationalised network IRP run**

| | Existing Dealers | | | Rationalised network | | | Change | | | % Change | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Manufacturer | Sales | Dealers | Average | Sales | Dealers | Average | Sales | Dealers | Average | Sales | Dealers | Average |
| Ford (01) | 16.615 | 105 | 158 | 16.893 | 70 | 241 | 277 | -35 | 83 | 1,7 | -33,3 | 52,5 |
| Audi-VW (03) | 19.745 | 101 | 195 | 19.478 | 59 | 330 | -267 | -42 | 135 | -1,4 | -41,6 | 68,9 |
| GM/Opel (09) | 17.590 | 100 | 176 | 17.676 | 60 | 295 | 86 | -40 | 119 | 0,5 | -40,0 | 67,5 |
| Mercedes (17) | 697 | 45 | 15 | 916 | 19 | 48 | 218 | -26 | 33 | 31,3 | -57,8 | 211,0 |
| Nissan (19) | 6.152 | 84 | 73 | 6.943 | 50 | 139 | 792 | -34 | 66 | 12,9 | -40,5 | 89,6 |
| Rover (23) | 1.275 | 23 | 55 | 2.078 | 9 | 231 | 803 | -14 | 175 | 63,0 | -60,9 | 316,5 |

**Figure 4 – Pre and post-rationalised network of dealers in Denmark**




## 3. Applications of models in business: ii) microsimulation models

### 3.1. Introduction

Microsimulation involves the estimation of individual or household data using chain conditional probabilities. The use of conditional probabilities allows the incorporation of the widest range of available known data to reconstruct detailed micro-level populations. This approach is necessary because the only source of small area population data, the Census, is released as a series of separate, pre-determined, tabulations such as age by sex by marital status for individuals and tenure by socio-economic group of head of household for households. By turning each tabulation into a probability conditional upon an attribute from another table, the data contained in separate tables may be linked together. An example is useful at this stage: for a given area, the age, sex, marital status distribution of household heads, P(A,S,M), and the distribution of age of household head by tenure type, P(A,T), are typically known. What is unknown is the distribution of head of households by age, sex, marital status **and** tenure type. In order to connect the information from the two distributions, the probability of household tenure type given age of head of household, Pr(T|A), is calculated. Each head of household of known age, sex and marital status is then probabilistically assigned a tenure type, given his/her age. Totalling the results of this process after every head of household has been allocated a tenure type, the estimated distribution of head of households by age, sex, marital status and tenure type, P(A,S,M,T) for the area under study may then be found. Note that, because tenure type allocation is based on random sampling (also known as Monte Carlo simulation), the estimate for P(A,S,M,T) obtained will vary slightly as the procedure is repeated. However, this random variation is effectively eliminated if the average of a number of runs (normally five) is taken.

Thus, the most important output from the microsimulation process is a data set which gives an estimate of 'missing data' – either new combination of variables not available from the Census or new data which is created by the fusion of a number of different data sets. The estimation of household income is a good example of the latter (Birkin and Clarke 1989, Ballas and Clarke 2001). However, once constructed, the models provide another powerful set of techniques for forecasting and prediction – i.e what are the impacts of change sin any of the variables used in the models. The use of the model will be illustrated below using two case studies.

### 3.2 Water demand

This section presents a summary of the findings of a research project commissioned by Yorkshire Water, aimed at investigating the feasibility of producing micro-level population and water demand estimates in order to help inform strategic, planning and other management decision making processes. An important aspect of policy underpinning this research work has been that of the impending (enforced) change in the basis of customer billing. Water companies in the UK have traditionally based the majority of their billing on the rateable value of the properties in their area. This has been recognised as an inappropriate methodology following the termination of regular surveys to update rateable values and water companies are now having to plan for an alternative means of billing customers.

At the same time efficient use of water has also emerged as a major policy issue, with water industry regulators in particular expressing concern over this matter. Further emphasis on efficient use of water has resulted from an increased focus on sustainability, an issue that has entered the public and environmental arena since the start of this project. One approach to the efficient use of water is through conservation based on pricing as a tool to promote reduced use of water. Demand management has seldom been addressed in this way in the water industry, which to date has typically equated supply with demand. In addition to influencing the efficiency of water use by customers through pricing strategy, efficiency of water use also needs to be tackled through the direct control of leakage, particularly since amount of leakage is an issue that influences both public opinion and National Rivers Authority policy. To this end, the identification of leakage and the prioritisation of leakage control strategies is an area of policy of great importance to all water plcs. The use of microsimulation techniques in water demand estimation has potential to aid the identification of high and low leakage areas, thus providing a basis for rational economic appraisal of leakage control measures by Yorkshire Water.

The methodological approach adopted during the course of the project was to estimate individually the characteristics of every household in the city of Leeds. Six household and individual characteristics derived from the 1991 Census were incorporated into the Leeds micro-level population estimation model. These are household location (ward), household tenure and property types, number of rooms, number of household members and socio-economic group of the head of household. As an example, the representation of a single household taken from the all Leeds population estimate may be pictured as:

| Household | Location | Tenure type | Property type | Number of Rooms |
|---|---|---|---|---|
| 1 | Cookridge, Leeds | Owner-occ. | Semi-detached | Seven |

| Individual | Socio-economic group |
|---|---|
| 1 | Professional worker – self-employed |
| 2 | |
| 3 | |
| ... | |
| … | |
| N | |

where N is the number of persons in the household.

Each of the characteristics included are known to be direct or indirect determinants of household water demand. To calibrate the model, data on household water consumption were supplied by Yorkshire Water in the form of 4039 metered households selected from across Yorkshire. Analysis of these data allowed the quantification of the effect various household attributes have on household water demand. From this analysis it was found that a combination of five household attributes appeared to best explain inter-household variations in water demand. These are, in order of importance, number of occupants, presence/absence of a dishwasher, presence/absence of a washing machine, number of bedrooms and household type (detached, semi, terraced, flat or bungalow). Some examples of the variation in water demand by such variables is shown in table 4.

**Table 4 – Tabulations of average household water demand ($m^3$) from a 513 household sample of metered households in Yorkshire by:**

| ***property type*** | | ***number of household residents*** | | ***number of bedrooms in housing unit*** | |
|---|---|---|---|---|---|
| Bungalow | 87.82 | 1 | 50.18 | 0 (bedsitter) | 18.51 |
| Detached | 134.92 | 2 | 104.92 | 1 | 49.73 |
| Flat/Maisonette | 61.17 | 3 | 136.22 | 2 | 72.52 |
| Semi-detached | 104.83 | 4 | 172.61 | 3 | 106.39 |
| Through-terrace | 102.32 | 5 | 187.11 | 4 | 148.30 |
| | | 6+ | 416.64 | 5 | 167.66 |
| | | | | 6+ | 251.90 |

In addition to generating household conditions such as tenure and number of occupants for each household in Leeds, we also needed to estimate household ownership rates of washing machines and dishwashers. The relationship between these five household attributes and predicted water demand may be expressed mathematically as follows:

$$W = 0.0296\,k + 0.33\,OC + 0.257\,DW + 0.109\,WR + 0.0536\,BD - 0.0287\,HT$$

where W = water demand
k = constant
OC = number of household occupants
DW = presence/absence of a dishwasher in household
WR = presence/absence of a washing machine in household
BD = number of bedrooms in household
HT = household property type

**Figure 5 – Spatial distributions of the best estimate of annual household water demand in Leeds**

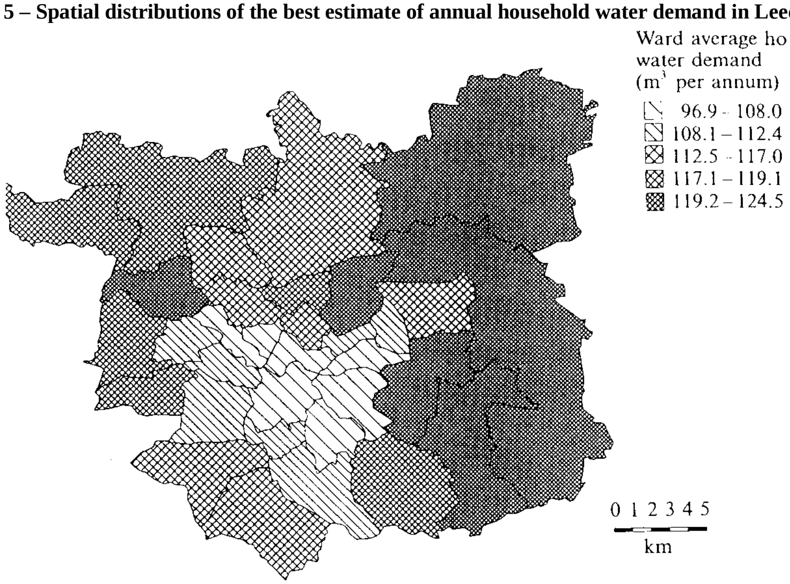



Figure 5 presents the spatial distributions of the best estimate of annual household water demand in Leeds. To recap, these estimates have been made using Monte Carlo sampling from the conditional probability of water demand given household property type, number of occupants, number of bedrooms and presence/absence of a washing machine and dishwasher estimated in the previous Section. Figure 5 reveals that wards on the Eastern fringe of Leeds do indeed seem to have the highest average annual household water demand levels, although one ward, Horsforth, on the Western edge of Leeds also falls into this category. Inner city wards are also shown as being those where the lowest average levels of household water demand are to be expected. Average household water demand then increases away from the city centre (except to the East) in more or less concentric circles. For more details of this applications ee Williamson et al (1996), Williamson

(2001).

### 3.3. Rural policy change in Ireland

The argument in this section is drawn from Ballas et al (2001). There is considerable interest currently in the European Union (EU) in rural policy. The EU Commission and Member States have, over recent years, placed particular emphasis on rural development with special reference to:

- enhancing the competitiveness of rural areas, maintaining and creating employment
- reducing socio-economic disparities between regions, adapting to new market place realities
- providing appropriate training and converting and re -orientating agricultural production potential (CEC, 1988, Ray, 1998).

This concern for rural development was encapsulated in The Cork Declaration (European Conference on Rural Development, 1996) which announced a 10 point rural development Programme for the European Union. It asserted that sustainable rural development must be put at the top of the agenda of the European Union and defined its aims as reversing rural out-migration, combating poverty, stimulating employment and equality of opportunity, and responding to growing requests for more quality, health, safety, personal development and leisure, and improved rural well-being. It also asserted that a rural development policy must be multi-disciplinary in concept, and multi-sectoral in application, with a clear territorial dimension.

With increasing recognition that rural development is not synonymous with agricultural development (and with an increasing range and diversity of policy measures) there is a need to develop tools of analysis which will enable the impact of rural development policy to be assessed ex post and also to enable the potential impact of new policies to be assessed before implementation. The Rural Economy Research Centre (RERC) of Ireland, Teagasc (a state sponsored research and development organisation, which had already established credibility in projecting the impact of changes in agricultural policy) was eager to embark on similar analyses in relation to changes in rural development policy. Accordingly a programme of collaboration between the Rural Economy Research Centre and the University of Leeds was emaciated with a view to developing a model, which would be capable of analysing the differential impacts of changing rural development. The overall aim of this project is to develop a new framework for the analysis of the Irish rural economy based on spatial microsimulation. Table 5 summarises the different stages of the project, whereas table 6 outlines the data sets that can be used as input into the spatial microsimulation modelling exercise.

**Table 5 – Spatial microsimulation tasks**

| | |
|---|---|
| **Task 1** | Building 1991 demographic and labour force model (static microsimulation) |
| **Task 2** | Projecting population and labour force for 1996 (dynamic microsimulation) |
| **Task 3** | Projecting population and labour force for 2002 (dynamic microsimulation) |

**Table 6 – Data sets that can be utilised in a spatial microsimulation framework**

| Data Set | Spatial Scale | Source |
|---|---|---|
| 1991 & 1996 Census data | DED | CSO |
| Digital Boundaries | DED | OS |
| Migration flows | County | CSO |
| Labour Force Survey | National | CSO |
| Household Budget Survey | National | CSO |
| ESRI Survey of Income Distribution, Poverty and Usage of State Services | National | ESRI |
| Vital Statistics | County | CSO |

As can be seen in table 5, the first stage of the project involves the building of a static spatial microsimulation demographic and labour force model for the Irish rural economy. The task of this model is to provide reliable estimates of the spatial distribution of the Irish population in 1991. In particular, the task of the model is to produce spatially disaggregated population microdata at the District Electoral Division level for all the rural areas of Ireland. The methodologies that can be employed to achieve this were described in the previous section.

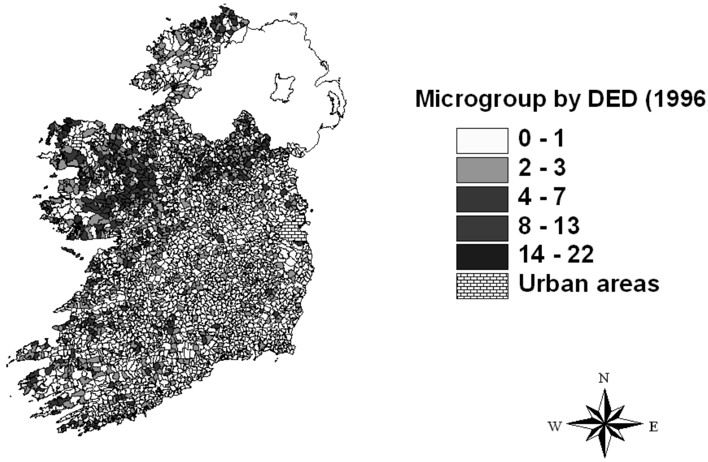
One of the particular strengths of spatial microsimulation is its policy-relevancy (see for instance Ballas and Clarke, 2001). In a rural policy context, spatial microsimulation models can be used to identify particular socio-economic groups that may be affected by agricultural policy reforms, such as the CAP reforms. Matthews (2000) points out that the number of farms in Ireland has been falling over time and that most of this decline is concentrated among smaller farms whose number fell from 85,000 to 67,000 between 1992 and 1999, while the number of larger farms remained stable at around 75,000. Further, Matthews (2000) distinguishes between three groups of farmers:

1. Full-time commercial farmers (low-cost efficient production of commodities – requiring scale and specialisation – dairy farmers, intensive enterprises, large-scale tillage and livestock). This type of farmers comprises 30% of the total
2. Mixed-income farmers (off-farm jobs) comprising 30% of the total
3. Marginal/transitional (elderly, relying on social welfare payments for a significant part of their income – farms non viable, no potential inheritors living on their farms) comprising 40% of the total.

The above analysis provided by Matthews can be very useful in the context of rural policy design and analysis. However, it would also be useful to know the geographical location of these particular population groups (and this is not publicly available). One of the advantages of spatial microsimulation modelling is that it can be used to identify the size and spatial location of particular population groups like the above. For instance, figure 6 depicts the estimated spatial distribution of male farmers in 1996, who were older than 55 years old and had a farm with size less than

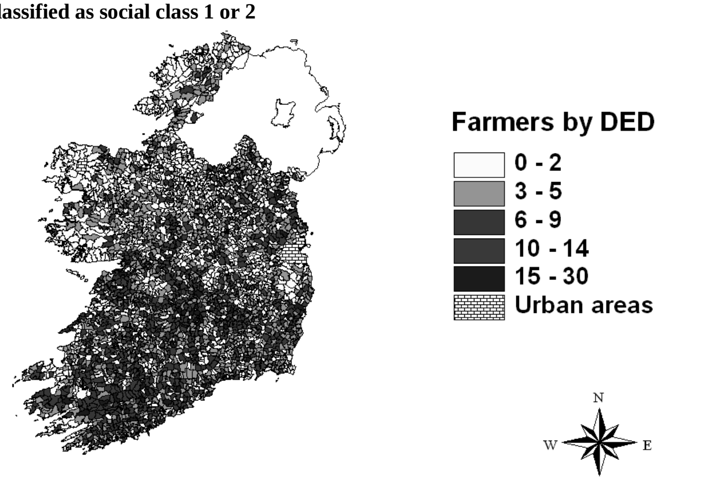
30 acres[1]. It can be argued that many of these farmers can be placed between the third category (marginal/transitional) described above.

**Figure 6 – Estimated spatial distribution of farmers belonging to type 3 (1996, Irish DEDs)**




The European Commission has calculated that 80% of the support provided by the CAP goes to the largest 20% of farmers. Using a microsimulation model it is possible to provide estimates of CAP-related support distributions at the small area level. Figure 7 depicts the spatial distribution of full-time farmers who own large farms and are classified as social class 1 or 2. It can be assumed that these farmers may have the capacity to have relatively high production and attract large amounts of CAP-related support.

**Figure 7 – Estimated spatial distribution of full-time farmers who own large farms and are classified as social class 1 or 2**




[1] These estimates were based on conditional probability data calculated from the 1991 and 1996 Irish Census SAPS

## 4. Conclusions

It is undoubtedly the case that the experience of building models for the private sector has increased our knowledge greatly on the properties of our techniques. However, we argue that models are best used within a process that involves other components. Using models to reproduce existing data is a start but this alone is not enough. Interpretation of the processes is the key. Our modelling techniques should be seen as a set of methods and tools, which are context dependent. Models should be built around client problems not the other way round. Generalized, off-the-shelf, modelling packages are unlikely to address the critical issues in each and every unique market. This argument holds for modelling routines which are part of more generic GIS solutions (see Benoit and Clarke 1997 for more detailed arguments here). Every aspect of model design and calibration is context dependent. As well as modelling skills, the analyst must offer the client sector knowledge, ingenuity, imagination and problem-solving skills. Fortunately, there are many people in regional science that possess these strengths!

## References:


Bailly A., Coffey W. (1994) Regional science in crisis: a plea for a more open and relevant approach, Papers in Regional Science, 73, 3-14

Ballas, D. and Clarke G. P. (2001), Modelling the local impacts of national social policies: a spatial microsimulation approach, Environment and Planning C: Government and Policy 19, pp. 587 – 606

Ballas D., Clarke G.P., Commins P. (2001) Spatial microsimulation for rural policy analysis, paper presented at the 41st European Regional Science Association Congress, Zagreb, August.

Benoit, D., Clarke G.P. (1997) Assessing GIS for retail location planning, Journal of Retail and Consumer Services, 4(4), 239-258

Birkin M., Clarke G.P., Clarke M., Wilson A.G. (1996) Intelligent GIS: location decisions and strategic planning, Cambridge: Geoinformation.

Birkin M, Clarke M, (1989), The generation of individual and household incomes at the small area level, Regional Studies 23 535-548

Birkin M., Culf R. (2001) Optimal distribution strategies, in G.P.Clarke & M.Madden (eds) Regional science in business, Springer, Berlin, 223-242

CEC (Commission of the European Communities) (1988). The Future of Rural Society, Supplement 4/88, Bulletin of the European Commission, Luxembourg

Clarke G.P., Clarke M. (2001) Applied spatil interaction modelling, in G.P.Clarke & M.Madden (eds) Regional science in business, Springer, Berlin, 137-158

Clarke G.P., Langley R., Cardwell W. (1998) Empirical applications of dynamic spatial interaction models, Computers, Environment and Urban Systems, 22(2), 157-184

Clarke G.P., Madden M. (eds, 2001) Regional science in business, Springer, Berlin.

Clarke G.P., Stillwell J.C.H. (eds, 2003) Applied GIS and spatial analysis, Wiley, Chichester.

European Conference on Rural Development (1996), The Cork Declaration -a Living Countryside, Cork, November

Eyre H., Clarke G.P., Rees P.H. (2000) Retail interaction models in practice: refining the attractiveness term, working paper, School of Geography, University of Leeds.

Fischer M., Getis A. (1999) New advances in spatial interaction theory, Papers in Regional Science, 78(2), 117-118

Fotheringham A.S. (1986) Modelling hierarchical destination choice, Environment and Planning A, 18, 401-418

Harris B., Wilson A.G. (1978) Equilibrium values and dynamics of attractiveness terms in production constrained spatial interaction models, Environment and Planning A, 10, 371-388

Matthews, A (2000), Agriculture, Food Safety and Rural Development, in J W O'Hagan (ed.), The Economy of Ireland: policy & performance of a European Region, Gill & Macmillan, Dublin, pp. 232-259

Pacione M. (1974) Measures of the attraction factor, Area, 6, 279-282

Ray, C. (1998). Culture, intellectual property and territorial rural development, Sociologia Ruralis, 15(3) pp 257-267.

Spencer A.H. (1978) Deriving measures of attractiveness for shopping centres, Regional Studies, 12, 713-726

Williamson P. (2001) An applied microsimulation model, in G.P.Clarke & M.Madden (eds) Regional science in business, Springer, Berlin, 243-268

Williamson, P., Clarke, G.P., McDonald, A.T. (1996), Estimating small area demands for water with the use of microsimulation, in G.P. Clarke (ed.), Microsimulation for Urban and Regional Policy Analysis, Pion, London, pp. 117-148

Wilson A.G. (1970) Entropy in urban and regional modelling, London: Pion.

Wilson A.G. (1974) Urban and regional models in geography and planning, Chichester: Wiley.